Executive Steering Committee
For A.C.E. Policy II
(ESCAP II)
Report 17
October 17, 2001

ESCAP II: Census 2000 Housing Unit Coverage Study

Diane F. Barrett,
Michael Beaghen,
Damon Smith, and
Joseph Burcham

Decennial Statistical
Studies Division and
Planning, Research, and
Evaluation Division

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

CONTENTS

EXECUTIVE SUMMARY

What is the Housing Unit Coverage Study?

The Housing Unit Coverage Study measures the Census 2000 housing unit coverage using data from the Accuracy and Coverage Evaluation. It uses dual system estimation to estimate the net coverage of housing units enumerated in the 2000 Census. The study also examines the percent of housing units in the population sample not matched to the census (P-Sample nonmatches) and the percent of housing units erroneously included in the enumeration sample (erroneous enumerations). These two components of the dual system estimate, evaluated separately, are used to measure the completeness and accuracy of the final address list containing all housing units existing in the United States on April 1, 2000. Understanding housing unit coverage is essential to evaluating coverage measurement procedures.

How good was the national coverage of housing units in the 2000 Census? How did it compare to 1990?

Coverage of housing units enumerated in the 2000 Census was comparable to the housing unit coverage in 1990. Both censuses had an undercount of less than 1.0 percent. Table 1 shows:

- The net undercount of housing units in the 2000 Census was 0.61 percent which was not significantly different than the net undercount in 1990 at 0.96 percent.
- For occupied housing units, no significant difference was observed between the 2000 and 1990 coverage. The net undercount was 0.33 percent in 2000 and 0.53 percent in 1990.
- The net undercount for vacants was 3.37 percent in 2000 which was not significantly different from the 4.71 percent net undercount in 1990.

Table 1 National Percent Net Undercount (Standard error)

Status	2000 A.C.E.	1990 HUCS
National	**0.61 (0.16)**	**0.96 (0.24)**
Occupied	0.33 (0.13)	0.53 (0.21)
Vacant	3.37 (0.98)	4.71 (1.26)

How did the coverage of occupied housing units compare to 1990?

The coverage for occupied housing units was consistent with what we found in 1990 for various research categories such as occupancy status, tenure, and type of enumeration area. The results that support this finding are:

- In 2000, vacant housing units (3.37 percent) were significantly undercounted more than occupied units (0.33 percent) which results in a difference of 3.04 percentage points. In 1990, the difference between the net undercount for vacants (4.71percent) and occupied units (0.53 percent) was 4.18 percentage points.

- The Accuracy and Coverage Evaluation has evidence that the census misclassified a small number of vacant units as occupied. Only a small percentage of this misclassification was attributable to unclassified imputations.

- As in 1990, the 2000 coverage for non-owners was not significantly different than for owners. The net undercount for owners was 0.12 percent in 2000 and 0.37 percent in 1990. The net undercount for non-owners was 0.57 percent in 2000 and 0.80 percent in 1990.

- The size of the metropolitan statistical area had no impact on the coverage of housing units in mailout/mailback areas. For occupied housing units, there were no significant differences between the net undercounts for mailout/mailback areas in small (0.11 percent), medium (0.30 percent), or large (0.53 percent) metropolitan areas versus all other types of enumeration areas (0.22 percent).

For occupied housing units, were there any coverage results that were unexpected?

Yes, for two characteristics: (1) Housing units with Non-Hispanic Black householders had a lower coverage rate than housing units with Non-Hispanic White or Some Other Race householders. (2) Occupied small multiunits were actually overcounted. Using the Bonferroni multiple comparison tests, we found the following:

> *Between occupied housing units with Non-Hispanic Black householders and those with Non Hispanic White or Some Other Race householders*:

- Even though the estimated net undercount of -0.45 percent for housing units with Non-Hispanic Black householders was not significantly different from zero, it was significantly lower than the estimated net undercount of 0.38 percent for housing units with Non-Hispanic White or Some Other Race householders.

- However, their nonmatched percents were not significantly different. The percent of P-Sample nonmatches was 2.34 percent for housing units with Non-Hispanic Black householders and 2.56 percent for housing units with Non-Hispanic White or Some Other Race householders.

- Nor were their percents of erroneous enumerations significantly different. Housing units with Non-Hispanic Black householders were erroneously enumerated at 1.87 percent in the census, which was not significantly different than the 1.37 percent erroneous enumerations for housing units with Non-Hispanic White or Some Other Race householders.

Among occupied small multiunits with 2 to 9 housing units at the basic street address (small multiunits), occupied large multiunits with 10 or more housing units at the basic street address (large multiunits) and occupied single units.

- Small multiunits had a net undercount of -1.30 percent which was significantly different from zero. Small multiunits were overcounted in 2000 but were significantly undercounted in 1990 at 2.11 percent.

- The overcount for small multiunits (-1.30 percent net undercount) was also significantly different than the coverage for single units (0.62 percent net undercount) but not significantly different from large multiunits. Large multiunits had a net undercount of -0.08 percent which was not significantly different from zero.

- Addresses for small multiunits were the most problematic among the three types of structures for the census. The percent of P-sample nonmatches (4.98 percent) and the percent of erroneous enumerations (3.74 percent) for small multiunits were both significantly higher than for single units (2.32 percent and 1.09 percent, respectively) and for large multiunits (2.39 percent and 1.89 percent, respectively).

What were the reasons for erroneous enumeration of occupied housing units?

The major reason for erroneous enumeration of an occupied housing unit in 2000 was that the address was not a housing unit; that is, it was nonresidential or did not exist on Census day. We did not distinguish between those addresses that were nonresidential (that is, group quarters, commercial, uninhabitable, and so on) or nonexistent (such as vacant lots, demolished, burned down, unable to locate and so on). These have been combined into one type of erroneous enumeration category as "not a housing unit."

- Over half (57.05 percent) of all erroneous enumerations were not housing units. Of the occupied erroneous enumerations, the highest percentage (45.27 percent) was attributed to the "not a housing unit" category.

- In 1990, not a housing unit (37.3 percent) and duplicates (33.4 percent) were both major reasons. However, duplicates had the highest percentage (40.7 percent) of occupied erroneous enumerations. For 2000, the duplicate percentage for occupied erroneous enumerations was lower at 28.69 percent.

- Duplicates in both the 1990 and 2000 censuses accounted for a large portion of the erroneous enumerations. Even though there were more duplicates in 1990 than in 2000, the proportion of duplicates for 2000 may be understated. The percentage of duplicates did not include late census adds (reinstatements). It is likely that some of the reinstatements may actually have been duplicates.

What implications do these results have on the adjustment decision?

The results of housing unit coverage do not bear directly on the question of adjustment. It appears that we did not find any unusual results that would indicate problems in the housing unit component of the A.C.E.

1. BACKGROUND

1.1 What questions does this report answer?

The report provides answers to the following questions for the 2000 Census:

- What was the net coverage of housing units? How did it compare to 1990?
- Did the census misclassify vacant housing units as occupied?
- What was the coverage of housing units by various research categories (occupancy status, tenure, type of structure and so on)? How did it compare to 1990?
- What was the major reason for erroneous enumerations? How did it compare to 1990?

1.2 What is the Housing Unit Coverage Study?

The Housing Unit Coverage Study (HUCS) is a study that measures the Census 2000 housing unit coverage using data from the Accuracy and Coverage Evaluation (A.C.E.). It uses dual system estimation to estimate a net undercount of housing units enumerated in the 2000 Census. The study also examines the percent of housing units in the population sample not matched to the census (P-Sample nonmatches) and the percent of housing units erroneously included in the enumeration sample (erroneous enumerations). These two components of the dual system estimate, evaluated separately, are used to measure the completeness and accuracy of the final address list containing all housing units existing in the United States on April 1, 2000.

1.3 What was the national coverage of housing units in 2000? How did it compare to 1990?

The national coverage of housing units for 2000 was comparable to the housing unit coverage in 1990. Both censuses had a net undercount of less than 1.0 percent. Table 1 shows:

- The net undercount of housing units in the 2000 Census was 0.61 percent which was not significantly different than the net undercount in 1990 at 0.96 percent.

- For occupied housing units, no significant difference was observed between the 2000 and 1990 coverage. The net undercount was 0.33 percent in 2000 and 0.53 percent in 1990.

- The net undercount for vacants was 3.37 percent in 2000 which was not significantly different from the 4.71 percent in 1990.

Table 1 National Percent Net Undercount (Standard error)

Status	2000 A.C.E.	1990 HUCS
National	**0.61 (0.16)**	**0.96 (0.24)**
Occupied	0.33 (0.13)	0.53 (0.21)
Vacant	3.37 (0.98)	4.71 (1.26)

2. METHODS

2.1 Research Categories

Tables 2 through 8 provide the percent of P-sample nonmatches, percent of erroneous enumerations, percent of late census adds (reinstatements) and the percent net undercount by the following research categories:

- occupancy status,
- tenure,
- race/Hispanic origin of the householder,
- type of structure and
- Metropolitan Statistical Area/Type of Enumeration (MSA/TEA) group.

2.2 Production Dual System Estimates (DSEs) versus Single Cell DSEs

The tables in this report provide net coverage estimates using the single cell DSE. Where obtainable, the production DSEs are also provided.

- **Production DSEs** - For some of the research categories we have obtained production dual system estimates (DSEs) by summing over appropriate poststrata. We cannot obtain production DSEs for some research categories that were combined in one or more poststrata (such as type of structure and MSA/TEA) or where the research category was not a poststratum variable (such as tenure).

- **Single cell DSEs** - For research categories where a production DSE is not obtainable, we calculated the net coverage estimate using a single cell DSE within each category. This method does not take the post stratification into account and thus the estimated undercounts may be understated.

2.3 Percent Net Undercount Comparisons to 1990

Where comparable, the percent net undercount from the 1990 Housing Unit Coverage Study (HUCS) is provided in the tables for the various research categories as documented in Childers (1993).

2.4 Significance Testing

We used the Bonferroni multiple comparisons test to compare coverage estimates between various characteristics.

3. LIMITS

The data in this report are not the official final numbers. Although the data in this report are not absolutely final, we believe the rates will not change in a material way, thus the numbers can be used for ESCAPII decisions**.**

4. RESULTS

4.1 What were the housing unit coverage estimates by occupancy status?

Coverage for both occupied and vacant units was not significantly different than in 1990. See Table 2:

- In 2000, vacant housing units (3.37 percent) were significantly undercounted more than occupied units (0.33 percent) which resulted in a difference of 3.04 percentage points. In 1990, the difference between the net undercount for vacants (4.71 percent) and occupied units (0.53 percent) was 4.18 percentage points.

- Both the percent of nonmatches (13.54) and the percent of erroneous enumeration (10.50) for vacant units are high.

Table 2 Housing Unit Coverage Estimates by Occupancy Status (Standard error)

Status	Percent P-Sample Nonmatches	Percent Erroneous Enumeration	Percent Late Census Adds	Net Percent Undercount		
				2000 single cell DSE	2000 production DSE	1990 HUCS
Occupied	2.61 (0.11)	1.51 (0.07)	0.86	0.27 (0.13)	0.33 (0.13)	0.53 (0.21)
Vacant	13.54 (0.79)	10.50 (0.67)	1.03	2.40 (0.99)	3.37 (0.98)	4.71 (1.26)
National	**3.62 (0.15)**	**2.31 (0.11)**	**0.87**	**0.48 (0.17)**	**0.61 (0.16)**	**0.96 (0.24)**

4.2 Did the 2000 Census misclassify vacant housing units as occupied?

Yes, there is evidence from the A.C.E. that the 2000 Census misclassified a small number of vacant units as occupied. However, only a small percentage of this misclassification was attributable to unclassified imputations of persons in vacant units.

It is important to note that misclassification does not affect the overall undercoverage of housing units but more than likely had only a minimal effect on the net percent undercount of vacants as well as the net percent undercount of occupied housing units. While a net undercount asserts the census failed to enumerate vacant units, misclassification asserts the census enumerated the vacant unit but counted it as an occupied unit.

We used the same occupancy statuses that were used for the A.C.E. postratification. That is, the final status from the Hundred Percent Census Unedited File (HCUF) and the final estimation outcome code derived from the results from the A.C.E person interviewing. Further research is necessary to evaluate the extent to which the A.C.E. classification was correct.

- Among matched E-sample housing units, there were 2.84 million weighted census housing units classified as occupied that the A.C.E. classified as vacant (see Appendix A, Table A-1). The census on the other hand classified 1.2 million weighted units as vacant that the A.C.E. classified as

occupied. Thus there was a net misclassification of about 1.6 million vacant housing units as occupied units (about 1.6 percent of the matched E-sample units).

- Of these 2.84 million units that the census classified as occupied but the A.C.E. classified as vacant, about 164,359 (about 6 percent) had only non-data defined persons, and of these, 62,008 (0.38 percent) were attributable to unclassified imputations (see Appendix A, Table A-2). Thus we could rule out unclassified imputation as a major source of classification error of occupancy status.

4.3 What were the housing unit coverage estimates by tenure?

We compared the net percent undercount of occupied units between owner and non-owner. Tenure was not one of the poststrata variables for producing housing unit dual system estimates**,** thus Table 3 below shows the net percent undercount using the single cell dual system estimates.

- As in 1990, the 2000 coverage for non-owners was not significantly different than for owners. The net undercount for owners was 0.12 percent in 2000 and 0.37 percent in 1990. The net undercount for non-owners was 0.57 percent in 2000 and 0.80 in 1990.

Table 3 Housing Unit Coverage Estimates by Tenure (Standard Error)

Tenure	Percent P-Sample Nonmatches	Percent Erroneous Enumeration	Percent Late Census Adds	Percent Net Undercount		
				2000 single cell DSE	2000 production DSE	1990 HUCS
Owner	2.14 (0.11)	1.26 (0.07)	0.77	0.12 (0.13)	na	0.37 (0.21)
Non-owner	3.56 (0.22)	2.02 (0.15)	1.02	0.57 (0.26)	na	0.80 (0.39)
Vacant	13.54 (0.79)	10.50 (0.67)	1.03	2.40 (0.99)	3.37 (0.98)	4.71 (1.26)
National	**3.62 (0.15)**	**2.31 (0.11)**	**0.87**	**0.48 (0.17)**	**0.61 (0.16)**	**0.96 (0.24)**

na-not available

4.4 What were the housing unit coverage estimates by race/Hispanic origin of householder?

We analyzed the coverage of housing units by race/Hispanic origin of the householder. The race/Hispanic origin groupings or domains were defined during person DSE processing. For housing unit DSE processing, occupied housing units were classified by the domain of the householder (person1). Refer to the Appendix, Table A-3 for the percent of E-Sample housing units each domain represents. See Table 4 on the next page.

Housing units with Non-Hispanic Black householders had a lower coverage rate than housing units with Non-Hispanic Whites or Some Other Race householders.

- Even though the estimated net undercount of -0.45 percent for housing units with Non-Hispanic Black householders was not significantly different from zero, it was significantly lower than the estimated net undrcount of 0.38 percent for housing units with Non-Hispanic White or Some Other Race householders.

- However, their nonmatched percents were not significantly different. The percent of P-sample nonmatches was 2.34 percent for housing units with Non-Hispanic Black householders and 2.56 percent for housing units with Non-Hispanic Whites or Some Other Race householders.

- Nor were their percents of erroneous enumerations significantly different. Housing units with Non-Hispanic Black householders were erroneously enumerated at 1.87 percent in the census, which was not significantly different than the 1.37 percent erroneous enumerations for housing units with Non-Hispanic White or Some Other Race householders.

Another finding which relates to housing units with Hispanic householders and housing units with Non-Hispanic White and Some Other Race householders was:

- The coverage was not significantly different. Housing units with Hispanic householders were undercounted at 0.06 percent and housing units with Non-Hispanic Whites or Some Other Race householders were undercounted at 0.38 percent.

Table 4 Housing Unit Coverage Estimates by Race/Hispanic Origin of Householder in Occupied Units (Standard Error)

Race/Hispanic Origin of Householder	Percent P-Sample Nonmatches	Percent Erroneous Enumeration	Percent Late Census Adds	Percent Net Undercount		
				2000 single cell DSE	2000 production DSE	1990 HUCs
Non-Hispanic White or "Some other race"	2.56 (0.12)	1.37 (0.07)	0.83	0.38 (0.14)	na	na
Non-Hispanic Black	2.34 (0.22)	1.87 (0.20)	0.93	-0.45 (0.29)	-0.44 (0.29)	na
Hispanic	3.01 (0.29)	1.98 (0.19)	1.00	0.06 (0.35)	0.19 (0.35)	na
Non-Hispanic Asian	3.00 (0.51)	2.09 (0.34)	0.69	0.26 (0.62)	0.22 (0.61)	na
Native Hawaiian or Pacific Islander	7.11 (2.54)	1.34 (0.53)	0.99	4.91 (2.62)	5.67 (2.82)	na
American Indian or Alaska Native- on reservation	6.64 (1.36)	3.79 (0.68)	1.22	1.78 (1.44)	1.88 (1.47)	na
American Indian or Alaska Native - off reservation	3.93 (0.95)	2.45 (0.44)	1.24	0.30 (1.00)	na	na
Total Occupied	**2.61 (0.11)**	**1.51 (0.07)**	**0.86**	**0.27 (0.13)**	**0.33 (0.13)**	**0.53 (0.21)**

4.5 What were the housing unit coverage estimates by type of structure?

We examined coverage estimates by the three type of structure categories defined for 2000. In 1990 there were five type of structure categories. The number of units at the basic street address was used as a proxy for type of structure. Refer to Tables 5 and 6 for the following findings:

Among occupied small multiunits with 2 to 9 housing units at the basic street address (small multiunits), occupied large multiunits with 10 or more housing units at the basic street address (large multiunits) and occupied single units.

- Small multiunits had a net undercount of -1.30 percent which was significantly different from zero. Small multiunits were overcounted in 2000 but were significantly undercounted in 1990 at 2.11 percent.

- The overcount for small multiunits (-1.30 percent net undercount) was also significantly different than the coverage for single units (0.62 percent net undercount) but not significantly different from large multiunits. Large multiunits had a net undercount of -0.08 percent which was not significantly different from zero.

- Addresses for small multiunits were the most problematic among the three types of structures for the Census. The percent of P-sample nonmatches (4.98 percent) and the percent of erroneous enumerations (3.74 percent) for small multiunits were both significantly higher than for single units (2.32 percent and 1.09 percent, respectively) and for large multiunits (2.39 percent and 1.89 percent, respectively).

Table 5 Housing Unit Coverage Estimates by Type of Structure - Total (Standard Error)

Type of Structure	Percent P-Sample Nonmatches	Percent Erroneous Enumeration	Percent Late Census Adds	Percent Net Undercount		
				2000 single cell DSE	2000 production DSE	1990[1] HUCS
Single Units	3.18 (0.15)	1.78 (0.07)	0.68	0.76[2] (0.16)	na	na
Small Multiunits 2 to 9 HUs	6.94 (0.57)	4.78 (0.23)	2.48	-0.17 (0.64)	na	2.25 (0.65)
Large Multiunits 10 or more HUs	3.39 (0.44)	2.97 (0.51)	0.57	-0.13 (0.54)	na	na
National	**3.62 (0.15)**	**2.31 (0.11)**	**0.87**	**0.48 (0.17)**	**0.61 (0.16)**	**0.96 (0.24)**

Table 6 Housing Unit Coverage Estimates by Type of Structure - Occupied (Standard Error)

Type of Structure	Percent P-Sample Nonmatches	Percent Erroneous Enumeration	Percent Late Census Adds	Percent Net Undercount		
				2000 single cell DSE	2000 production DSE	1990[1] HUCS
Single Units	2.32 (0.12)	1.09 (0.06)	0.64	0.62[2] (0.13)	0.63 (0.13)	na
Small Multiunits 2 to 9 HUs	4.98 (0.43)	3.74 (0.20)	2.63	-1.30 (0.48)	na	2.11 (0.59)
Large Multiunits 10 or more HUS	2.39 (0.31)	1.89 (0.32)	0.60	-0.08 (0.44)	na	na
Total Occupied	**2.61 (0.11)**	**1.51 (0.07)**	**0.86**	**0.27 (0.13)**	**0.33 (0.13)**	**0.53 (0.21)**

na-not available

[1] 1990 HUCS Types of structure.

	Total	Occupied
Single (no mobile homes)	0.76 (0.23)	0.05 (0.18)
Medium multiunits (10-49 HUs)	-2.41 (1.22)	-2.19 (1.12)
Large multiunits (50+ hus)	-0.94 (1.23)	0.09 (0.52)
Other (mostly mobile homes)	4.46 (1.28)	4.50 (1.26)

[2] 2000 ACE single category includes mobile homes.

4.6 What were the coverage estimates by Metropolitan Statistical Area/Type of Enumeration Area (MSA/TEA) Group?

We have calculated coverage estimates by MSA/TEA groups which were defined during housing unit post stratification. It appears that the size of the metropolitan statistical area had no impact on coverage in mailout/mailback areas. See Tables 7 and 8 for the following comparisons:

- For all housing units as well as for occupied housing units, there was no significant difference between the net undercounts for mailout/mailback areas in small, medium, or large versus all other types of enumeration areas.

- The net undercount for non-mailout/mailback areas shrunk to almost nothing (0.22 percent) when we looked at occupied units. This may be attributed to the large percent of late adds in this category.

Table 7 Housing Unit Coverage Estimates by Metropolitan Statistical Area/Type of Enumeration Area (MSA/TEA) Group - Total

MSA/TEA	Percent P-Sample Nonmatches	Percent Erroneous Enumeration	Percent Late Census Adds	Percent Net Undercount		
				2000 single cell DSE	2000 production DSE	1990 HUCS
Large MSA Mailout/Mailback	3.01 (0.24)	2.13 (0.17)	0.69	0.22 (0.29)	na	na
Medium MSA Mailout/Mailback	2.41 (0.22)	1.60 (0.14)	0.42	0.41 (0.25)	na	na
Small MSA&NonMSA Mailout/Mailback	3.59 (0.34)	2.62 (0.38)	0.42	0.58 (0.35)	na	na
All Other TEAs	6.52 (0.44)	3.38 (0.14)	2.31	1.01 (0.47)	na	na
Total	**3.62 (0.15)**	**2.31 (0.11)**	**0.87**	**0.48 (0.17)**	**0.61 (0.16)**	**0.96 (0.24)**

na - not available

Table 8 Housing Unit Coverage Estimates by Metropolitan Statistical Area/Type of Enumeration Area (MSA/TEA) Group - Occupied Units

MSA/TEA	Percent P-Sample Nonmatches	Percent Erroneous Enumeration	Percent Late Census Adds	Percent Net Undercount		
				2000 single cell DSE	2000 production DSE	1990 HUCS
Large MSA Mailout/Mailback	2.36 (0.20)	1.57 (0.14)	0.71	0.11 (0.24)	na	na
Medium MSA Mailout/Mailback	1.85 (0.18)	1.14 (0.12)	0.43	0.30 (0.21)	na	na
Small MSA&NonMSA Mailout/Mailback	2.52 (0.21)	1.56 (0.15)	0.45	0.53 (0.26)	na	na
All Other TEAs	4.44 (0.35)	2.01 (0.11)	2.31	0.22 (0.37)	na	na
Total Occupied	**2.61 (0.11)**	**1.51 (0.07)**	**0.86**	**0.27 (0.13)**	**0.33 (0.13)**	**0.53 (0.21)**

na-not available

4.7 What were the reasons for erroneous enumerations of housing units?

The major reason for erroneous enumeration of an occupied housing units in 2000 was that the address was not a housing unit; that is, it was nonresidential or did not exist on Census day. We did not distinguish between those addresses that were nonresidential (that is, group quarters, commercial, uninhabitable, and so on) or nonexistent (such as vacant lots, demolished, burned down, unable to locate, and so on). These have been combined into one type of erroneous enumeration category as "not a housing unit." See Tables 9 and 10 for the following results:

- Over half (57.05 percent) of all erroneous enumerations were not housing units. Of the occupied erroneous enumerations, the highest percentage (45.27 percent) was attributed to the "not a housing unit" category.

- In 1990, not a housing unit (37.3 percent) and duplicates (33.4 percent) were both major reasons. However, duplicates had the highest percentage of occupied erroneous enumerations. For 2000, the duplicate percentage for occupied erroneous enumerations was lower at 28.69 percent.

- Duplicates in both the 1990 and 2000 censuses accounted for a large portion of the erroneous enumerations. Even though there were more duplicates in 1990 than in 2000, the proportion of duplicates for 2000 may be understated. The percentage of duplicates did not include late census adds (reinstatements). It is likely that some of the reinstatements may actually have been duplicates.

Table 9 Percent of Erroneous Enumeration by Reason (Standard Error)

	Total		Occupied	
Reason	**2000 Percent**	**1990 Percent**	**2000 Percent**	**1990 Percent**
Duplicates	24.81 (2.76)	33.4 (na)[3]	28.69 (1.29)	40.7 (na) [3]
Geocoding errors	16.15 (1.72)	16.2 (3.0)	23.67 (1.60)	22.3 (4.0)
Not a housing unit	57.05 (2.51)	37.3 (3.4)	45.27 (1.51)	24.4 (2.7)
Unresolved	1.99 (0.56)	2.8 (0.4)	2.37 (0.55)	2.0 (0.3)
Insufficient Information	na	10.2 (2.0)	na	10.6 (2.1)
Total	**100.0**	**99.9**	**100.0**	**100.0**

na-not available

Table 10 Percent of E-Sample by Type of Erroneous Enumeration for Census 2000 (Standard Error)

Reason	Total	Occupied
Duplicates	0.57 (0.80)	0.43 (0.04)
Geocoding errors	0.37 (0.04)	0.36 (0.04)
Not a housing unit	1.32 (0.06)	0.68 (0.04)
Unresolved	0.05 (0.01)	0.04 (0.01)
Total	**2.31**	**1.51**

5. CONCLUSION

[3] In 1990, percents and standard errors were calculated separately for within block and surrounding block. The percents have been combined for comparisons. The 1990 standard error was not recalculated.

Coverage of housing units in the 2000 Census was good, at least as good as it was in 1990. For some characteristics the net coverage improved. Changes to census procedures such as development of the address list and reinstating late census adds may have had an impact in keeping the percent net undercounts low.

The results of housing unit coverage do not bear directly on the question of adjustment. It appears that we did not find any unusual results that would indicate problems in the housing unit component of the A.C.E.

6. REFERENCES

Burcham, Joseph (2001) *"Accuracy and Coverage Evaluation: Creation of Input Files for Computing the Housing Unit Dual System Estimates"* DSSD Census 2000 Procedures and Operations Memorandum Series Updated Q-55, Bureau of the Census, June 26, 2001.

Childers, Danny (1993) *"Coverage of Housing in the 1990 Decennial Census"* 1990 Decennial Census Preliminary Research and Evaluation Memorandum Series No. 253, Bureau of the Census, October 7, 1993.

Childers, Danny (2000) *"Accuracy and Coverage Evaluation: The Design Document"*, DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-1, Bureau of the Census, January 26, 2001.

Griffin, Richard and Haines, Dawn (2000) *"Accuracy and Coverage Evaluation Survey: Post-stratification for Dual System Estimation,"* DSSD Census 2000 Procedures and Operations Memorandum Series Q-21, Bureau of the Census, January 12, 2000.

Hefter, Steven P. (2001) *"Accuracy and Coverage Evaluation: Housing Unit Dual System Estimation Programming Specifications(U.S.)*, DSSD Census 2000 Procedures and Operations Memorandum Series Reissue Q-56, Bureau of the Census, August 14, 2001.

APPENDIX A. Additional Tables

Table A-1 Census versus A.C.E. Occupancy Status for Matched E-Sample Housing Units

	A.C.E.			
Census	Non-Interview	Occupied	Vacant	Total
Occupied	2,108,057	90,309,469	2,840,794	95,258,320
	2.06%	88.26%	2.78%	93.10%
Vacant	348,934	1,201,095	5,511,784	7,061,814
	0.34%	1.17%	5.39%	6.90%
Total	**2,456,991**	**91,510,564**	**8,352,578**	**102,320,134**
	2.40%	**89.44%**	**8.16%**	**100.00%**

Table A-2. Source of Final Status for the Vacant Housing Units with Non-data Defined People

Source of Final Status	**Weighted Housing Units**
Respondent-initiated Return	4,504
Enumerator Completed Form	97,846
Unclassified Imputation	62,008
Mail Return Checkin only	42,317
Field Status and Pop Count	4,994
Occupied Field Status but no Pop Count	14,697
Total	**164,359**

Table A-3 Percent of E-Sample by Research Category

Research Category	Characteristic	Percent
Occupancy Status	Occupied	91.06
	Vacant	8.94
Tenure	Owner-Occupied	60.76
	Non-owner-Occupied	30.30
Race/Hispanic Origin of Householder	Non-Hispanic White or "Some other race"	69.44
	Non-Hispanic Black	10.28
	Hispanic	7.92
	Non-Hispanic Asian	2.72
	Native Hawaiian or Pacific Islander	0.13
	American Indian or Alaska Native- on resv	0.11
	American Indian or Alaska Native -off resv	0.46
Type of Structure	Single	72.94
	Multiunit with 2-9 HUs	11.55
	Multiunit with 10+ HUs	15.51
	Single-Occupied	67.13
	Multiunit with 2-9 HUs-Occupied	10.24
	Multiunit with 10+ HUs-Occupied	13.68
Metropolitan Statistical Area/Type of Enumeration Area (MSA/TEA)	Large MSA MO/MB (Mailout/mailback)	28.01
	Medium MSA MO/MB	31.04
	Small MSA&NonMSA MO/MB	21.41
	All Other TEAs	19.54
	Large MSA MO/MB-Occupied	26.53
	Medium MSA MO/MB-Occupied	28.91
	Small MSA&NonMSA MO/MB-Occupied	19.33
	All Other TEAs-Occupied	16.28

APPENDIX B. Technical Documentation

B.1 Housing unit dual system estimation output files

For the detailed file specifications and record layouts of the housing unit dual system estimation output files, see:

- Burcham, Joseph (2001) *"Accuracy and Coverage Evaluation: Creation of Input Files for Computing the Housing Unit Dual System Estimates"* Update to Q-55, (update 1 dtd 6/26/01)

- Hefter, Steven P. (2001) *"Accuracy and Coverage Evaluation: Housing Unit Dual System Estimation Programming Specifications(U.S.)*, Reissue of Q-56, (revised draft dtd 814/01*)*

B.2 Variable Recodes and New Variables

B.2.1 For P-Sample Processing (PHUFO_US)

OCCUP (Occupancy Status)

If FINOUTC in (1, 2, 3) then OCCUP = 1
Else if FINOUTC in (10, 11) then OCCUP = 2
Else OCCUP = 3

B.2.2 For E-Sample Processing (EHUFO_US)

OCCUP (Occupancy Status)

If FINST = 1 then OCCUP = 1
Else if FINST = 2 then OCCUP = 2
Else if FINST = 3 then do
    If NP = 0 then OCCUP = 2
    Else if NP ne 0 then OCCUP = 1

TENURE

If TENURE in (1, 2) then TENURE = 1
If TENURE in (3, 4) then TENURE = 2

ENUMTYPE

If DUP2 > 0 and FHICODE NOT = UE then ENUMTYPE = DE
Else if FHICODE = GU then ENUMTYPE = GE
Else if FHICODE in (P,MU,UE) then ENUMTYPE = UE
Else ENUMTYPE = FHICODE

B.3 Calculation of Percent P-sample Nonmatch

Filename: PHUFO_US.DAT
Variables: (PRHU, PRM, TRIMWTP, TESWGT)

Formula : Nonmatch rate = $1 - \frac{M}{Np}$ or $\frac{NMp}{Np}$

where:

M = the weighted number of P-sample matched housing units, or

$$M = \sum_{k} PRM * TRIMWTP * TESWGT$$

NM_p =the weighted number of P-sample nonmatched housing units, or

$$NM_p = \sum_{k} (1 - PRM) * TRIMWTP * TESWGT$$

N_p = the weighted number of P-sample housing units, or

$$N_p = \sum_{k} TRIMWTP * TESWGT$$

k = the subset of housing units of interest; i.e., vacant, single unit, etc.

Percent P-sample nonmatch = nonmatch rate * 100

B.4 Calculation of Percent Erroneous Enumeration

Filename: EHUFO_US
Variables: (PRCE, DUPFACT, TRIMWTE, TESWGT, TES2WGT)

Formula: Erroneous Enumeration Rate = $1 - \frac{CE}{N_e}$ or $\frac{EE}{N_e}$

where:

CE = the weighted estimate of the number of correct enumerations in the E-sample, or

$$CE = \sum_{k} PRCE2 * TRIMWTE * TESWGT * TES2WGT$$

Where PRCE2 = PRCE*DUPFACT

EE = the weighted estimate of the number of erroneous enumerations in the E-sample, or

$$EE = \sum_{k} (1 - PRCE2) * TRIMWTE * TESWGT * TES2WGT$$

N_e = the weighted number of E-sample housing units, or

$$N_e = \sum_{k} TRIMWTE * TESWGT * TES2WGT$$

Percent of Erroneous Enumerations = Erroneous Enumeration rate * 100

*Note when calculating Erroneous Enumeration rates one may save processing time by only processing records where ESAMP = 1.

B.5 Dual System Estimates

Filename1: CHUFO_US.DAT
Variables: (TRIMWTE, FINST) for variables UBSA2, TENURE, MSATEA, DOMAIN, and REGION

Filename2: POST_US
Variables: (CCWO) for the variable OCCUP

Formula: The formula for the dual system estimate of the population of HUs is:

$$DSE = \frac{(C)\ (\frac{CE}{N_e})}{\frac{M}{N_p}}$$

where CE, N_e, M, N_p are defined as above and:

C = the count of housing units in the census (*does not include late census adds*) or for variables UBSA2, TENURE, MSATEA, DOMAIN and REGION, from CHUFO_US

$$C = \sum_k IND(FINST \neq 3)$$

Where IND(statement) = 1 if the statement is true, 0 otherwise.

or, for variable OCCUP, from POST_US

$$C = \sum_k CCWO$$

B.6 Determining late census adds

Filename1: CHUFO_US.DAT
Variables: (TRIMWTE, FINST) for variables UBSA2, TENURE, MSATEA, DOMAIN, and REGION

Filename2: POST_US
Variables: (CCWO) for the variable OCCUP

Formula: For UBSA2, TENURE, MSATEA, DOMAIN and REGION, from CHUFO_US

$$\text{Late Census Adds} = \frac{\sum_k \text{IND}(\text{FINST} = 3)}{C}$$

Where C is defined above.

For OCCUP, from POST_US (CCWO)

$$\text{Late Census Adds} = \frac{\sum_k \text{CCW} - \text{CCWO}}{\sum_k \text{CCWO}}$$

Percent of late census adds = Late Census Adds * 100

B.7 Calculation of Net Percent Undercount

Filename1: CHUFO_US.DAT
Variables: (TRIMWTE, FINST) for variables UBSA2, TENURE, MSATEA, DOMAIN, and REGION

Filename2: POST_US
Variables: (CCWO) for the variable OCCUP

Formula**:** Undercount rate = DSE minus the census count including late adds, divided by the DSE, or

$$\text{Undercount rate} = \frac{\text{DSE} - C^*}{\text{DSE}}$$

where DSE is defined in B.5 above and:

C^* = the count of housing units in the census (*includes late census adds*) or for variables UBSA2, TENURE, MSATEA, DOMAIN and REGION, from CHUF_US

$$C^* = \sum_k 1$$

or for variable OCCUP, from POST_US

$$C^* = \sum_k \text{CCW}$$

Percent Undercount = Undercount rate * 100

B.7 Type of Erroneous Enumeration

Filename: EHUFO_US.DAT
Variables: DUP2, FHICODE

Formulas: Percentage of Erroneous Enumerations with ENUMTYPE of 'EE' =

$$\frac{ee\_sum}{EE}$$

Percentage of Erroneous Enumerations with ENUMTYPE of 'GE' =

$$\frac{ge\_sum}{EE}$$

Percentage of Erroneous Enumerations with ENUMTYPE of 'de' =

$$\frac{de\_sum}{EE}$$

Percentage of Erroneous Enumerations with ENUMTYPE of 'ue' =

$$\frac{ue\_sum}{EE}$$

where:

ee_sum = nationwide weighted estimate of the records with ENUMTYPE = 'EE', or

$$\text{ee\_sum} = \sum_{ee} (1 - PRCE2) * TRIMWTE * TESWGT * TES2WGT$$

ge_sum = nationwide weighted estimate of the records with ENUMTYPE = 'GE', or

$$\text{ge\_sum} = \sum_{ge} (1 - PRCE2) * TRIMWTE * TESWGT * TES2WGT$$

de_sum = nationwide weighted estimate of the records with ENUMTYPE = ‘DE’, or

$$\text{de\_sum} = \sum_{de} (1 - PRCE2) * TRIMWTE * TESWGT * TES2WGT$$

ue_sum = nationwide weighted estimate of the records with ENUMTYPE = ‘UE’, or

$$\text{ue\_sum} = \sum_{ue} (1 - PRCE2) * TRIMWTE * TESWGT * TES2WGT$$

EE = nationwide weighted estimate of erroneous enumerations, or

$$\text{EE} = \sum (1 - PRCE2) * TRIMWTE * TESWGT * TES2WGT$$

PRCE2 = PRCE*DUPFACT

*Note that when calculating rates in this section, one may save processing time by only processing records where ESI = 1.

B.8 Identifying evidence of misclassification of vacant housing units

B.8.1 Create the dataset HUPER

Merge the E-sample HUDSE Input file, EHUFO, to the estimation file for census people EFINUS, by cluster and census ID (CLUST, CID)
In HUPER keep only the records where ESI = 1 and FHICODE = M
Drop records from EHUFO if there is no matching EFINUS record
From EFINUS keep only the variables CLUST, CID, and CEPROBF

B.8.2 Create the dataset MAX

Sort HUPER by CLUST, CID
Maintain the variable MAXCE from data step to data step (retain statement)
For the first person record in each CID bygroup set MAXCE = CEPROBF
For each subsequent person record in the CID bygroup if CEPROBF > MAXCE set MAXCE = CEPROBF
For the last person record in each CID bygroup output to MAX

B.8.3. Create the dataset MAXPER

Merge MAX with the reformatted CUF (only those records in the E-sample).
From the reformatted CUF keep the variables CLUST, CID, ESAMP, INPS, INP, FINST and SFINST
Output to MAXPER a record for every E-sample record that is on the reformatted CUF
If a record is on the reformatted CUF and in the E-sample, but not on MAX, assign an arbitrary value to MAXPER not between 0 and 1 (for example, 5)

B.8.4. Create the dataset VAC

Merge MAXPER with PHUFO_US by CLUST and FHICID from PHUFO_US to CLUST CID on MAXPER
Do not include in the merge any records on PHUFO_US with FINOUTC = 12
Output to VAC only those records for there is a CLUST, CID match from both files
Create variable ACEOCCUP
    If FINOUTC in (4, 6, 9) then ACEOCCUP = nonint
    If FINOUTC in (1, 2, 9) then ACEOCCUP = occupy
    If FINOUTC in (10, 11) then ACEOCCUP = vacant
Create variable CENVAC
    If MAXCE=0 then do;
        If INP = 0 then CENVAC = vacant
        If INP > 0 then CENVAC = EE and II

If MAXCE = 5 and INP = 0 then CENVAC = finst?
If MAXCE = 5 and INP > 0 then CENVAC = II only
If MAXCE > 0 and MAXCE < 1 then CENVAC = unresolved
If MAXCE = 1 then CENVAC = CE people

B.8.5. Generate Table A.C.E. versus Census Occupancy Status for Matched E-Sample Units

From the dataset VAC produce crosstabulations of FINST with ACEOCCUP. Weight by the P-sample housing unit weights from the PHUFO_US, WEIGHTP and TESWGT

B.8.6 Generate Table Source of Final Status for E-Sample Units with only Non-Data Defined People

From the dataset VAC produce a tabulation of SFINST for the units with CENVAC = II only. Weight by the P-sample housing unit weights from PHUFO_US, WEIGHTP and TESWGT